Improved characterization of Clematis based on new chloroplast microsatellite markers and nuclear ITS sequences

Liu, Zhigao

2018-12

Liu , Z & Korpelainen , H 2018 , ' Improved characterization of Clematis based on new chloroplast microsatellite markers and nuclear ITS sequences ' , Horticulture environment and biotechnology , vol. 59 , no. 6 , pp. 889-897 . https://doi.org/10.1007/s13580-018-0090-3

http://hdl.handle.net/10138/307841
https://doi.org/10.1007/s13580-018-0090-3

acceptedVersion

Improved Characterization of *Clematis* Based on New Chloroplast Microsatellite Markers and Nuclear ITS Sequences

Zhigao Liu[1], and Helena Korpelainen[2]*

[1]*College of Landscape Architecture, Zhejiang A & F University, Lin'an, Zhejiang 311300, China*

[2]*Department of Agricultural Sciences, Viikki Plant Science Centre, PO Box 27, FI-00014 University of Helsinki, Finland*

*Corresponding author: helena.korpelainen@helsinki.fi

Abstract. Currently, there is a lack of genetic markers capable of effectively detecting polymorphisms in *Clematis*. Therefore, we developed new markers to investigate inter- and intraspecific diversity in *Clematis*. Based on the complete chloroplast genome of *Clematis terniflora*, simple sequence repeats were explored and primer pairs were designed for all ten adequate repeat regions (cpSSRs), which were tested in 43 individuals of 11 *Clematis* species. In addition, the nuclear ITS region was sequenced in 11 *Clematis* species. Seven cpSSR loci were found to be polymorphic in the genus and serve as markers that can distinguish different species and be used in different genetic analyses, including cultivar identification to assist the breeding of new ornamental cultivars.

Additional key words: genetic markers; genetic variation; molecular identification; ornamental plants, genetic relationships

Introduction

Simple sequence repeats (SSR) or microsatellites are widely distributed throughout nuclear and cytoplasmic genomes in eukaryotes (Litt and Luty, 1989). Because of their highly polymorphic nature, codominant inheritance, ubiquitous abundance and rapid mutation rates, SSRs have become useful markers for genetic diversity and population genetic analyses (Morgante et al., 2002; Selkoe and Toonen, 2006). Furthermore, SSRs can be detected by standard PCR methods and can be transferable to related taxa (Chen et al., 2015). The chloroplast genome is widely used in plant taxonomic and systematic studies (Rajendrakumar et al., 2007; Tambarussi et al., 2009) because it is usually maternally inherited and slowly evolving, and has a low frequency of genetic recombination (Birky and Walsh, 1988) but a high frequency of microsatellite repeats (Bryan et al., 1999; Provan et al., 1999). For these reasons, chloroplast markers are especially useful in genetic diversity and population genetic structure analyses, phylogenetic and phylogeographic analyses, and in hybrid identification of plants.

There are about 300 species in the genus *Clematis* L., which makes it the largest genus in Ranunculaceae (Grey, 2000; Wang and Li, 2005). The genus consists of typically vigorous, woody, climbing vines that are mainly distributed in the temperate zone of the northern hemisphere (Hao et al., 2013). The genus *Clematis* is famous for its diverse flower shapes and colors (Roh and Song, 1997); with the hundreds of cultivars, *Clematis* is known as the “Queen of the Vines”. *Clematis* cultivars are also used for medicinal purposes because triterpenoid saponins, flavonoids, and many

other compounds are present in various species (He et al., 2011).

Previous studies on the classification and phylogeny on *Clematis* were mainly based on morphological traits; phyllotaxy, sepals, cotyledon, and calyces have been used to characterize different species of *Clematis* (Goodley, 1977; Tobe et al., 1980; Keener and Dennis, 1982; Essig, 1991). More recently, different molecular marker systems have been applied to studies of *Clematis*. Inter-simple sequence repeat markers (ISSR) have been used to fingerprint 32 vining cultivars and five non-vining *Clematis* species to assess their genetic relationships and cultivar identification (Nicole and Stan, 2005). Random amplified polymorphic DNA (RAPD) has been used to confirm *Clematis* hybrids (Tao et al., 2010), and sequences of chloroplast DNA such as the *atpB-rbcL* spacer region, *matK*, *trnK*, *trnL* intron, and *trnL-trnF* spacer region, as well as the nuclear *actin* I intron have been used for the analysis of phylogenetic relationships within the *Clematis* genus (Johansson and Jansen, 1993; Johansson, 1995; Miikeda et al., 1999, 2006; Slomba et al., 2004; Shuang et al., 2016). Internal transcribed spacer (ITS) sequences have been analyzed to provide molecular evidence for the current phylogeny of the genus, and also for the identification of medicinal *Clematis* species (Ming et al., 2011, Xie et al., 2011; Xiao et al., 2012).

Classifications based on morphology and molecular methods have produced diverging results. According to Osamu et al. (2006), the inconsistencies between previous classification systems and molecular analyses indicate that several characters, such as the presence of filament hairs and the position of the inflorescences on the shoot (upper or lower part of stems), are homoplasious and do not clarify the

phylogenetic relationships. However, leaf margin characters, which had not been previously used to characterize *Clematis*, were found to be useful in defining subgenera (Osamu et al., 2006).

Currently, there is a lack of SSR or cpSSR markers capable of effectively detecting polymorphisms and confirming cultivar identity in *Clematis*. To improve precision in genetic analyses on *Clematis*, we developed cpSSR markers to investigate inter- and intraspecific diversity among *Clematis* samples. Such markers will also be useful tools when searching for cytoplasm donors to breed new ornamental cultivars. Additionally, we compared relationships among *Clematis* species based on two types of genetic tools: cpSSR markers developed in this study and the common "phylogeny tool" ITS sequencing, to test whether the results from both methods are congruent or not.

Materials and Methods

Sampling and DNA extraction

We analyzed 43 accessions that represent 11 *Clematis* species (Table 1) that were collected from the Zhejiang province of China. All vouchers are deposited at the Zhejiang A & F University. Total genomic DNA was extracted from dry leaf tissue using the E.Z.N.A Plant DNA Mini Kit Spin Protocol (Omega Bio-tek, Inc.) according to the manufacturer's instructions.

Primer design

The SSR locations were searched from the complete chloroplast genome of *Clematis terniflora* (GenBank, accession KM652489.1) using the SSRIT tool (http://archive.gramene.org/db/markers/ssrtool). Ten adequate SSR regions were identified using the following criteria: length (at least ten copies for mononucleotide repeats and at least six copies for other types of SSRs) of the repeat sequence and good flanking primer sites. Then, primer pairs were designed for these regions using Primer Premier 5.0 (http://downloads.fyxm.net/download-now-Primer-Premier-Others-Home-&-Educatio n-101178.html) using following parameters: primer length of 20-25 bp, a PCR product size of 100-300 bp, annealing temperature between 50-65°C, and a GC content of 30-60% (Table 2). Primers used for ITS sequencing were ITS1 (5'-CTT GGT CAT TTA GAG GAA GTA A-3') and ITS4 (5'-TCC TCC GCT TAT TGA TAT GC-3') (Gardes and Bruns, 1993)

PCR amplification and data analysis

PCR reactions for the selected ten SSR regions were carried out in 20-μl volumes by mixing the following components: 11 μl ddH_2O, 2 μl 10 × buffer, 0.4 μl 10 mM dNTPs, 0.6 μl Dynazyme II DNA polymerase (Thermo Fisher Scientific, 2 U μl^{-1}), 2 μl genomic DNA (about 20 ng) and 2 μl both primers (5 pmol μl^{-1}). The forward primers were fluorescently labeled with FAM or HEX. The PCR reactions were

117 carried out with an initial denaturation for 45 s at 94° C, followed by 35 cycles of 30 s
118 at 94°C, 30 s at the cpSSR-specific annealing temperature (Table 2), 40 s of
119 elongation at 72°C, and with a final elongation at 72°C for 5 min. After amplification,
120 the PCR products were diluted 1:20-1:200 depending on the concentration (the final
121 concentration about 1 ng μl^{-1}) with Milli-Q water. The DNA fragments were analyzed
122 using a capillary electrophoresis system 3730 DNA Analyzer (Applied Biosystems).
123 The DNA fragment sizes were determined using Peak Scanner ver. 2.0 (Applied
124 Biosystems). The observed number of alleles (Na), effective number of alleles (Ne)
125 and polymorphism information content (PIC) were estimated with PowerMarker
126 V3.25 (Liu and Muse, 2005). Phylogenetic trees were constructed with PowerMarker
127 V3.25 using the UPGMA method based on genetic distances described by Nei (1983).
128 The PCR reactions preceding ITS sequencing were performed in a total volume
129 of 20 µl that contained 13 µl ddH_2O, 2µl 10 × buffer, 0.4 µl 10 mM dNTP mix, 0.6 µl
130 of Dynazyme II DNA polymerase (2 U μl^{-1}), 2 µl genomic DNA (about 20 ng), and 1
131 µl both primers (5 pmol μl^{-1}). The PCR cycle was similar to that used for SSR
132 genotyping, but the annealing temperature was 50° C. Amplification products were
133 run in a 1% agarose gel, and the DNA fragments were excised and purified prior to
134 sequencing using the E.Z.N.A. Gel Extraction Kit (Omega Bio-Tek). Purified DNA
135 samples were sequenced at Macrogen Inc. using the same primers utilized in the PCR
136 reactions preceding ITS sequencing. The ITS sequences were manually checked using
137 Chromas 2.5.0 (http://chromas.software.informer.com) and then aligned using Clustal
138 X (http://www.clustal.org). The aligned data were analyzed using Mega 6.0

(http://www.megasoftware.net) and genetic distances between species were determined using the Kimura-2 method (Kimura, 1980). Phylogenetic trees were drafted using the UPGMA method and the Kimura-2 parameter model in Mega 6.0.

Results and Discussion

Development and assessment of cpSSR primers for identifying *Clematis* species

Aside from DNA sequencing-based investigations, there are only a limited number of molecular marker studies of *Clematis*. These include the report of ISSR primers (Nicole and Stan 2005) and randomly amplified polymorphic DNA (RAPD; Tao et al., 2010); however, none of these previous investigations have utilized SSR markers. The present study developed cpSSR markers based on the available chloroplast genome of *C. terniflora* and successfully applied them to investigate intra- and interspecific polymorphisms in *Clematis*. The polymorphic cpSSR markers can distinguish different species from each other and can be used in cultivar identification as well.

All ten cpSSR markers gave clear amplification products and seven were polymorphic within the genus (Table 3). A total of 28 alleles were discovered in the ten cpSSR loci among the 11 representative *Clematis* species. All amplified fragments were around the predicted sizes (Table 2, Table 4), indicating that the detected polymorphism mainly arose from variation in the number of cpSSR repeats. All

161 markers amplified in all species, except for loci Clecp2 and Clecp7 in *C. lasiandra*,
162 which was probably due to the failure of the primers to anneal at those locations. The
163 allele numbers per locus ranged from one to seven within the genus (Table 3). The
164 average number of alleles (Na), effective number of alleles (Ne), and polymorphism
165 information content (PIC) were 2.8, 1.9, and 0.29, respectively, for the whole dataset
166 (Table 3). In previous cpSSR-based studies, the mean PIC values equaled 0.19 in
167 common bean cultivars (Ceylan et al., 2014), 0.21 in cowpea (Pan et al., 2014), 0.32
168 in *Gladiolus* cultivars (Singh et al., 2017) and 0.60 in cotton (Li et al., 2014). These
169 examples show that PIC values of cpSSR markers can vary considerably and that the
170 value detected in this study falls within the observed range of variation.
171 The UPGMA dendrogram was constructed using all cpSSR allele information and
172 samples from the same species clearly grouped together (Figure 1). Five main clusters
173 were generated: *C. lasiandra* (37, 38), *C. patens* subsp. *tientaiensis* (41), and *C.*
174 *henryi* (27-30) grouped in clusters I, II, and III, respectively. *C. terniflora* (40), *C.*
175 *chinensis* var. *anhweiensis* (42), *C. courtoisii* (43), and *C. hancockiana* (35, 36) were
176 closely connected in the dendrogram and grouped with *C. apiifolia* (1-14) to form
177 cluster IV. *C. brevicaudata* (31-34), *C. finetiana* (15-26), and *C. uncinata* (39)
178 grouped as a cluster V.
179
180 **ITS sequencing**
181 The sequenced ITS region covered the whole distance from the end of the 18S
182 rRNA gene to the beginning of the 26S rRNA gene (GenBank accession numbers

KY201178- KY201188). The length of the entire ITS region (ITS1+5.8S+ITS2) from the 11 *Clematis* species varied between 534-562 bp. The ITS1 and ITS2 regions varied from 156-180 and 218-224 bp, respectively. *C. courtoisii* and *C. hancockiana* had the longest and *C. henryi* the shortest ITS regions. Among all species, the length of the 5.8S rRNA varied between 158-159 bp. The average GC content was 61.6% for the entire ITS region, 60.3% for ITS1, 68.5% for ITS2, and 53.2% for the 5.8S rRNA region (Table 5). Thus, some length variation was detected for the ITS regions among the *Clematis* species.

The numbers of variable sites and parsimony-informative sites within the ITS region equaled 79 and 38, and accounted for 13.8% and 6.6%,of the sites within the entire ITS region, respectively. The numbers of variable sites and parsimony-informative sites within ITS1, ITS2, and 5.8S rRNA regions were 39, 33, 7, and 24, 11, 3, respectively. Thus, variation was abundant and included both SNPs and indels. The studied species were characterized by considerable divergence, with genetic distances varying between 0.008-0.073 (Table 6). The smallest distance was 0.008, which was found between *C. hancockiana* and *C. patens* subsp. *tientaiensis*, while *C. finetiana* and *C. courtoisii* were the most divergent species with a distance value of 0.073. The average pairwise genetic distance equaled 0.039.

When a dendrogram was constructed for the 11 *Clematis* species using the UPGMA method, we discovered four clusters (Figure 2). The first cluster (I) was composed of *C. lasiandra*, *C. apiifolia*, *C. henryi*, *C. brevicaudata*, and *C. uncinata.* *C. chinensis* var. *anhweiensis*; *C. terniflora* were grouped in cluster II; *C. courtoisii*, *C.*

hancockiana, and *C. patens* subsp. *tientaiensis* were grouped in cluster III; and *C. finetiana* was separated from the other species in cluster IV. These sequencing results correspond previous analyses quite well (Ming et al., 2011; Xie et al., 2011).

Molecular markers as characterization and phylogenetic tools

It is widely recognized that the traditional use of morphological traits for taxon identification has several limitations, which include the misidentification of a taxon due to the phenotypic plasticity of the traits studied, the existence of cryptic taxa or the applicability of certain morphological keys only for a particular life stage (Valentini et al., 2009). In *Clematis*, the presence of filament hairs and the position of the inflorescences on the shoot (upper or lower part of stems) are homoplasious and do not inform phylogenetic relationships (Osamu et al., 2006). It is worth noting that convergent evolution may also confuse the interpretation of morphological traits; for instance, in the family Brassicaceae, there has been prevalent convergent evolution of several traits through time (Huang et al., 2016).

The species *C. lasiandra*, *C. apiifolia*, *C. henryi*, and *C. brevicaudata* included in our study share a common character of serrate leaf margins, which is supposed to be an indication of a close evolutionary relationship. Also, their pairwise genetic distances based on ITS sequences were quite small (0.010-0.022). The character state of entire leaf margins is shared by the other seven investigated *Clematis* species and therefore, the ITS-based relationships were found to correspond morphologically. For the most part, our results on *Clematis* relationships based on ITS regions agreed with

227 the ITS sequence analyses presented by Ming et al. (2011) and Xie et al. (2011),
228 which were based on nuclear ITS and plastid data.

229 In the cpSSR-derived tree, the seven species with entire leaf margins were sorted
230 into three clades: *C. terniflora*, *C. chinensis* var. *anhweiensis*, *C. courtoisii* and *C.*
231 *hancockiana* clustered together in the subgroup of clade IV; *C. uncinata* and *C.*
232 *finetiana* were in the same subgroup of clade V; and *C. patens* subsp. *tientaiensis*
233 formed clade II. Among them, *C. chinensis* var. *anhweiensis* and *C. terniflora* had a
234 close relationship based on both cpSSR and ITS data. Also, *C. courtoisii* and *C.*
235 *hancockiana* belonged to the same group in both UPGMA trees. However, the
236 grouping schemes of the other seven species were quite different in cpSSR- and
237 ITS-based trees. For example, *C. lasiandra* and *C. apifolia* belonged to two separate
238 clusters in the cpSSR tree, while they were grouped into the same cluster in the
239 ITS-based tree. Thus, the trees produced with cpSSR markers and ITS sequences did
240 not correspond with each other, possibly because of lineage sorting or introgression
241 (Wendel and Doyle, 1998), the relatively narrow range of markers and sequences used
242 in this study, or the size homoplasy for cpSSR markers, which may limit the
243 phylogenetic power of cpSSRs (Wheeler et al., 2014). The main strength of SSRs and
244 comparable markers is rather in species and genotype identification instead of
245 revealing phylogenetic relationships.

246 To the best of our knowledge, this is the first report on the development and use
247 of any kind SSR markers in the genus *Clematis*. These markers can be used in further
248 studies on genetic diversity, population genetics and phylogeography of *Clematis*, as

249 well as assist in the breeding of new ornamental cultivars.

250

251 **Acknowledgments:** This work was supported by the Nikolai and Ljudmila
252 Borisoff Foundation and the Fund of Forestry Key Disciplines of ZheJiang Province
253 (China).

254

255

256 **Literature Cited**

257

258 **Birky CW Jr, Walsh JB** (1988) Effects of linkage on rates of molecular evolution.
259 Proc Natl Acad Sci USA 85:6414–6418
260 **Bryan GJ, Mcbicoll J, Ramsay G, Meyer RC, De Jong WS** (1999) Polymorphic
261 simple sequence repeat markers in chloroplast genomes of *Solanaceous* plants.
262 Theor Appl Gen 99:859–867
263 **Ceylan A, Öcal, N, MikailAkbulut M** (2014) Genetic diversity among the Turkish
264 common bean cultivars (*Phaseolus vulgaris* L.) as assessed by SRAP, POGP and
265 cpSSR markers. Biochem Syst Ecol 54:219-229
266 **Chen H, Liu L, Wang L, Wang S, Somta P, Cheng X** (2015) Development and
267 validation of EST-SSR markers from the transcriptome of adzuki bean (*Vigna*
268 *angularis*). PLOS One 10:1–14
269 **Essig FB** (1991) Seedling morphology in *Clematis* (Ranunculaceae) and its
270 taxonomic implications. Sida 14:377–390

Gardes M, Bruns TD (1993) ITS primers with enhanced specificity for basidiomycetes-application to the identification of mycorrhizae and rusts. Mol Ecol 2:113–118

Goodley EJ (1977) Imbricate sepals in *Clematis*. *N Zeal J Bot* 15:775–776

Hao DC, Gu XJ, Xiao PG, Peng Y (2013) Chemical and biological research of *Clematis* medicinal resources. Chin Sci Bull 58:1120–1129

He YX, Li L, Zhang K, Liu ZR (2011) Cytotoxic triterpene saponins from *Clematis mandshurica*. J Asian Nat Prod Res 13:1104–1109

Huang C-H, Sun R, Hu Y, Zeng L, Zhang N, Cai L, Zhang Q, Koch MA, Al-Shehbaz I, Edger PP, J. Pires JC, Tan D-Y, Zhong Y, Ma H (2016) Resolution of Brassicaceae phylogeny using nuclear genes uncovers nested radiations and supports convergent morphological evolution. Mol Biol Evol 33:394–412

Johansson JT, Jansen RK (1993) Chloroplast DNA variation and phylogeny of the Ranunculaceae. Plant Syst Evol 187:29–49

Johansson JT (1995) A revised chloroplast DNA phylogeny of the Ranunculaceae. Plant Syst Evol 9 (Suppl.):253–262

Keener CS, Dennis WM (1982) The subgeneric classification of *Clematis* (Ranunculaceae) in temperate North America north of Mexico. Taxon 31:37–44

Kimura M (1980) A simple method for estimating evolutionary rate base substitution comparative studies nucleotide sequences. J Mol Evol 16:111–120

Li P, Li Z, Liu H, Hua J (2014) Cytoplasmic diversity of the cotton genus as

293 revealed by chloroplast microsatellite markers. Genet Resour Crop Evol
294 61:107-119
295 **Litt M, Luty JA** (1989) A hypervariable microsatellite revealed by in vitro
296 amplification of a dinucleotide repeat within the cardiac muscle actin gene. Am J
297 Hum Gen 44:397–401
298 **Liu K, Muse SV** (2005) PowerMarker: an integrated analysis environment for
299 genetic marker analysis. Bioinformatics 21:2128–2129
300 **Miikeda O, Kita K, Handa T, Yukawa T** (2006) Phylogenetic relationships of
301 *Clematis* (Ranunculaceae) based on chloroplast and nuclear DNA sequences. Bot
302 J Linn Soc 152:153–168
303 **Miikeda O, Koga S, Handa T, Yukawa T** (1999) Subgeneric relationships in
304 *Clematis* (Ranunculaceae) by DNA sequences. *In* S Andrews, A Leslie, C
305 Alexander, eds, Taxonomy of Cultivated Plants. Third International Symposium.
306 Royal Botanic Gardens, London, UK, pp. 355–358
307 **Ming J, Qiao ZY, Rong RL** (2011) ITS sequence analysis of eight medicinal
308 plants in *Clematis* L. Chin Trad Herb Drugs 42:1802–1806
309 **Morgante M, Hanafey M, Powell W** (2002) Microsatellites are preferentially
310 associated with nonrepetitive DNA in plant genomes. Nature Gen 30:94–200
311 **Nicole G, Stan CH** (2005) Intersimple sequence repeat fingerprinting and genetic
312 variation in a collection of *Clematis* cultivars and commercial germplasm. Hortic
313 Sci 40:1982–1987
314 **Pan L, Li Y, Guo R, Wu H, Hu Z, Chen C** (2014) Development of 12 chloroplast

315 microsatellite markers in *Vigna unguiculata* (Fabaceae) and amplification in
316 *Phaseolus vulgaris*. Appl Plant Sci 2: 1300075. DOI: 10.3732/apps.1300075
317 **Provan J, Soranzo N, Wilson NJ, Goldetein DB, Powell W** (1999) A low mutation
318 rate for chloroplast microsatellites. Genetics 153:943–947
319 **Rajendrakumar P, Biswal AK, Balachandran SM, Srinivasarao K, Sundaram**
320 **RM** (2007) Simple sequence repeats in organellar genomes of rice: frequency and
321 distribution in genic and intergenic regions. Bioinformatics 23:1–4
322 **Roh MD, Song CY** (1997) Effect of temperature and photoperiod on growth and
323 flowering of potted dwarf *Clematis*. Korean Soc Hort Sci 38:429–434
324 **Selkoe KA, Toonen RJ** (2006) Microsatellites for ecologists: a practical guide to
325 using and evaluating microsatellite markers. Ecol Lett 9:615–629
326 **Shuang XY, Hui JL, Le LL, Shuai L, Jin YL, Lin YP, Lei X** (2016) Taxonomic
327 status of *Clematis acerifolia* var. *elobata*, based on molecular evidence.
328 Phytotaxa 268:209–219
329 **Singh N, Pal AK, Roy RK, Tamta S, Rana TS** (2017) Development of cpSSR
330 markers for analysis of genetic diversity in *Gladiolus* cultivars. Plant Gene
331 10:31-36
332 **Slomba JM, Garey JR, Essig FB** (2004) The actin I intron–a phylogenetically
333 informative DNA region in *Clematis* (Ranunculaceae). SIDA 21:879–886
334 **Tambarussi EV, Melotto-Passarin DM, Gonzalez SG, Brigati JB, de Jesus FA,**
335 **Barbosa AL, Dressano K, Carrer H** (2009) In silico analysis of simple
336 sequence repeats from chloroplast genomes of Solanaceae species. Crop Breed

337 Appl Biotechn 9:344–352

338 **Tao Y, Yin WL, Mark SR** (2010) Confirmation of *Clematis* hybrids using

339 molecular markers. Sci Hort 125:136–145

340 **Tobe H** (1980) Morphological studies on the genus *Clematis* Linn., VII.

341 Reinvestigation of *Clematis* williamsii A. Gray and proposal of its taxonomic

342 transfer to *Clematopsis*. Bot Mag, Tokyo 93:135–148

343 **Valentini A, Pompanon F, Taberlet P** (2009) DNA barcoding for ecologists. Trends

344 in Ecol Evol 24:110-117

345 **Wang WT, Li LQ** (2005) A new system of classification of the genus *Clematis*

346 (Ranunculaceae). Acta Phytotax *Sin* 43:431–488

347 **Wendel JF, Doyle JJ** (1998) Phylogenetic incongruence: window into genome

348 history and molecular evolution. *In* D Soltis, P Soltis, J Doyle, eds, Molecular

349 Systematics of Plants, 2nd edn. Chapman and Hall, New York, pp. 265–296

350 **Xiao L, Pei W, Long Z, Hua AJ** (2012) rDNA-ITS sequence analysis of the

351 traditional mongolian medicinal plants in *Clematis* L. Acta Acad Med Neimongol

352 34:325–330

353 **Xie L, Wen J, Li LQ** (2011) Phylogenetic analyses of *Clematis* (Ranunculaceae)

354 Based on sequences of nuclear ribosomal ITS and three plastid regions. Syst Bot

355 36:907–921

Table 1. Information of the *Clematis* samples used for cpSSR and ITS analyses

Number	Taxon	Origin	Coordinates	Sample size	Altitude (m)
1	*C. finetiana*	Tian mu mountain (Lin an)	30°15’37.56’’N, 119°16’57.92’’E	3	499-589
2	*C. finetiana*				
3	*C. finetiana*				
4	*C. finetiana*	Wu li village (Lin an)	30°19’36.66’’N, 119°15’24.72’’E	5	399-500
5	*C. finetiana*				
6	*C. finetiana*				
7	*C. finetiana*				
8	*C. finetiana*				
9	*C. finetiana*	Ling long mountain (Lin an)	30°13’03.61’’N, 119°40’02.67’’E	6	111-144
10	*C. finetiana*				
11	*C. finetiana*				
12	*C. finetiana*				
13	*C. finetiana*				
14	*C. finetiana*				
15	*C. apiifolia*	Tian mu mountain (Lin an)	30°10’02.33’’N, 119°01’44.13’’E	9	533-610
16	*C. apiifolia*				
17	*C. apiifolia*				
18	*C. apiifolia*				
19	*C. apiifolia*				
20	*C. apiifolia*				
21	*C. apiifolia*				
22	*C. apiifolia*				
23	*C. apiifolia*				
24	*C. apiifolia*	Ban shan village (Lin an)	30°14’35.82’’N, 119°14’14.44’’E	3	553-661
25	*C. apiifolia*				
26	*C. apiifolia*				

27	*C. henryi*	Sankou mountain (Anji county)	30°24’30.05’’N, 119°40’35.96’’E	4	670-733
28	*C. henryi*				
29	*C. henryi*				
30	*C. henryi*				
31	*C. brevicaudata*	Longgang village (Shaoxing county)	29°50’17.47’’N, 120°39’26.93’’E	4	431
32	*C. brevicaudata*				
33	*C. brevicaudata*				
34	*C. brevicaudata*				
35	*C. hancockiana*	Qianqiu mountain (Lin an)	30°19’54.09’’N, 119°16’02.61’’E	2	579-597
36	*C. hancockiana*				
37	*C. lasiandra*	Wu li village (Lin an)	30°16’27.88’’N, 119°06’16.28’’E	2	503
38	*C. lasiandra*				
39	*C. uncinata*	Wu li village (Lin an)	30°16’27.88’’N, 119°06’16.28’’E	1	503
40	*C. terniflora*	Ling long mountain (Lin an)	30°13’18.21’’N, 119°40’06.62’’E	1	97
41	*C. patens* subsp. *tientaiensis*	Nan shan village (Tian tai county)	29°07’56.05’’N, 121°19’18.44’’E	1	735
42	*C. chinensis* var. *anhweiensis*	Liu an city anhui province	31°05’52.80’’N, 115°44’54.49’’E	1	791
43	*C. courtoisii*	Tian mu mountain (Lin an)	30°12’23.21’’N, 119°04’11.24’’E	1	477

Table 2. Ten cpSSR primer pairs developed for *Clematis* used in this study

Locus	Repeat type	Primer sequence (5' to 3')	TM (℃)	Product size (bp)
Clecp1	$[T/A]_{19}$	F: TTTGTTCATGCGGTACTCCTTT R: ATCTTGTCTATTCCCACGGTTC	59	138
Clecp2	$[T]_{14}$	F: AAGATACCGCTGTGCCAGGATA R: AGAAGCCGAGTAAGCGGATTGG	61	125
Clecp3	$[A/T]_{16}$	F: ATTTTCTATAACCTACCGTCTT R: TTGACTTCTACTATTTTGGTTG	50	116
Clecp4	$[A]_{12}$	F: GATAGGGGTCAATAAAAGAAAA R: ATAGGTGCATACAGTAGGCTCA	53	111
Clecp5	$[A/T]_{29}$	F: TTGTTTTCCACATCGTGATTTC R: TGTCCACTCACTTTATTTTCTGAAC	60	195
Clecp6	$[A/T]_{36}$	F: ATGGGGAGATAAAGAAATAGAG R: TACCAAAATAGGATGAAATAGG	52	152
Clecp7	$[G/A]_{22}$	F: ACCAGTTGTTGCTGATACCTCCTT R: CGGTCGTTGTGGTCGGACTCTA	61	128
Clecp8	$[A/T]_{21}$	F: AATGAAAGGGATGTTGAAAGAG R: CTGTCACGTACACGTAGGAATA	567	170
Clecp9	$[T/C]_{20}$	F: TAGGGATATGGAACGAAAGGAA R: ATTAATTCTCTAGCCCCGCTGT	60	204
Clecp10	$[T/A]_{30}$	F: TCTATGAAATGCCAATCCAACA R: AAAAACTTATAGGGCGTGGATAAA	56	209

Table 3. Summary of genetic variation statistics for each cpSSR locus among all *Clematis* samples

Locus	Samples	Na	Ne	PIC
Clecp1	43	7	3.23	0.65
Clecp2	43	1	1.00	0.00
Clecp3	43	1	1.00	0.00
Clecp4	43	4	2.88	0.59
Clecp5	43	2	1.05	0.04
Clecp6	43	2	1.37	0.24
Clecp7	43	1	1.00	0.00
Clecp8	43	3	1.62	0.35
Clecp9	43	2	1.93	0.37
Clecp10	43	5	3.74	0.68
Mean		2.8	1.9	0.29

Na, the observed number of alleles; Ne, effective number of alleles; PIC, polymorphism information content

Table 4. Detected allele sizes of 10 cpSSR loci in different *Clematis* species

Locus	*C. finetiana*	*C. apiifolia*	*C. henryi*	*C. brevicaudata*	*C. hancockiana*	*C. lasiandra*	*C. uncinata*	*C. terniflora*	*C.patens* subsp. *tientaiensis*	*Clematis chinensis var. anhweiensis*	*C. courtoisii*
Clecp1	140,141	135	133,136,137	135	127	135	133	135	127	135	127
Clecp2	122	122	122	122	122	-	122	122	122	122	122
Clecp3	112	112	112	112	112	112	112	112	112	112	112
Clecp4	108	106	105,106	107	106	105	106	106	106	106	107
Clecp5	193	193	193	193	193,194	193	193	193	193	193	193
Clecp6	149	149	149	150	149	150	149	149	150	149	149
Clecp7	124	124	124	124	124	-	124	124	124	124	124
Clecp8	168,169	168	173	168	168	169	169	168	169	168	168
Clecp9	204	203	203,204	203,204	204	204	203	204	204	204	204
Clecp10	231,232	229,230,231	232	229,230	232	230	231	232	237	232	232

-, no amplification

Table 5. GenBank accession numbers, length (bp), and GC content (%) of ITS sequences in 11 *Clematis* species

Accession number	Species	ITS region		ITS1		ITS2		5.8S	
		length	GC content	length	GC content	length	GC content	length	GC content
KY201178	*C. chinensis* var. *anhweiensis*	549	62.1	169	59.8	221	70.6	159	52.8
KY201179	*C. apiifolia*	552	61.4	173	60.1	220	68.2	159	53.5
KY201180	*C. brevicaudata*	543	61.3	164	59.2	220	68.6	159	53.5
KY201181	*C. courtoisii*	562	63.2	180	63.9	220	68.6	158	54.4
KY201182	*C. finetiana*	543	59.3	166	58.4	218	67.0	159	49.7
KY201183	*C. hancockiana*	562	63.4	180	63.9	224	69.2	158	54.4
KY201184	*C. henryi*	534	60.1	156	57.7	219	67.1	159	52.8
KY201185	*C. lasiandra*	550	60.4	172	59.9	219	66.2	159	52.8
KY201186	*C. patens* subsp. *tientaiensis*	560	63.4	180	64.4	222	68.9	158	54.4
KY201187	*C. terniflora*	549	61.8	169	59.2	221	70.1	159	52.8
KY201188	*C. uncinata*	541	60.8	163	57.1	219	69.0	159	53.5

Table 6. Pairwise divergence of *Clematis* species based on ITS sequences

	Species	Genetic distance										
		1	2	3	4	5	6	7	8	9	10	11
1:	*C. chinensis* var. *anhweiensis*	-										
2:	*C. terniflora*	0.020	-									
3:	*C. brevicaudata*	0.032	0.036	-								
4:	*C. henryi*	0.036	0.040	0.022	-							
5:	*C. lasiandra*	0.038	0.042	0.016	0.010	-						
6:	*C. apiifolia*	0.038	0.044	0.014	0.016	0.010	-					
7:	*C. finetiana*	0.060	0.065	0.042	0.044	0.038	0.040	-				
8:	*C. uncinata*	0.040	0.044	0.028	0.030	0.032	0.030	0.050	-			
9:	*C. courtoisii*	0.050	0.058	0.052	0.054	0.056	0.054	0.073	0.052	-		
10:	*C. hancockiana*	0.038	0.046	0.040	0.042	0.044	0.042	0.060	0.040	0.012	-	
11:	*C. patens* subsp. *tientaiensis*	0.036	0.044	0.036	0.038	0.040	0.038	0.056	0.038	0.020	0.008	-

Kimura-2 parameter model was used to calculate genetic distances

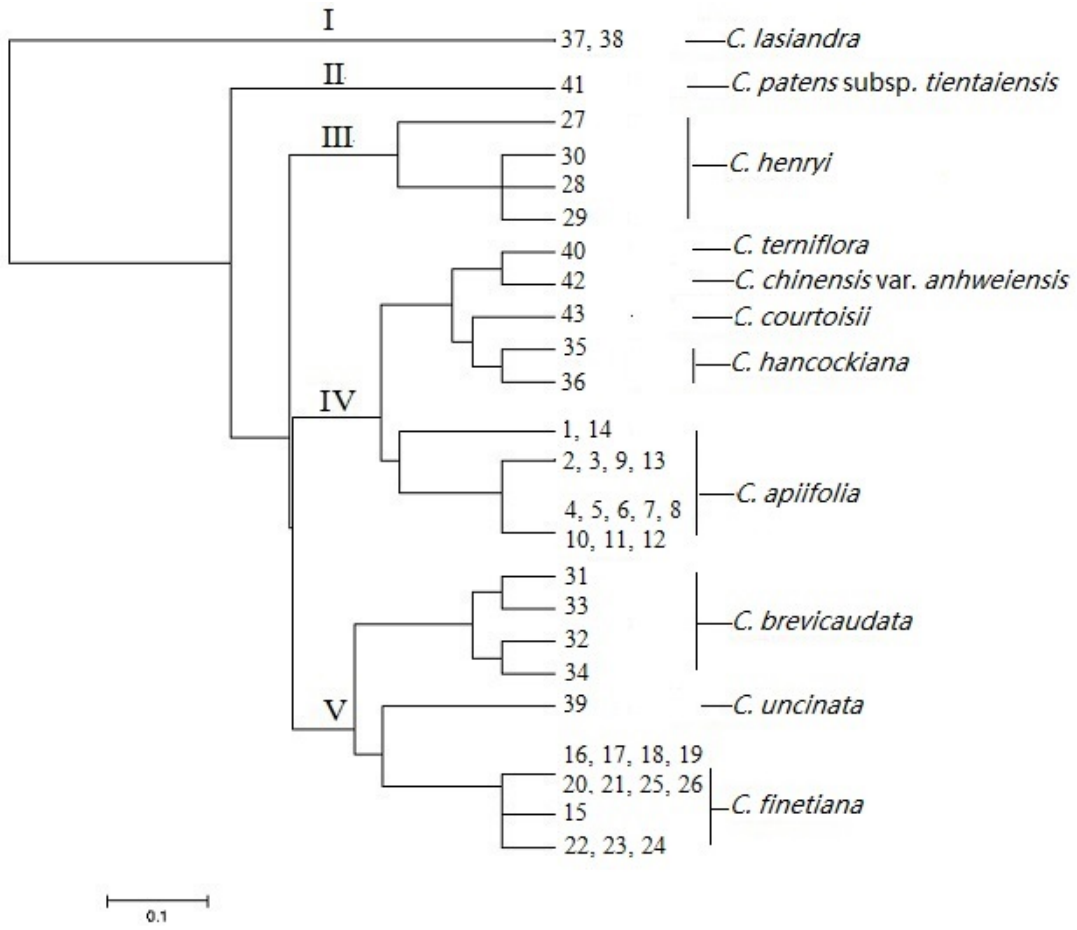

Fig. 1. A dendrogram of *Clematis* samples based on cpSSR variation. The sample numbers follow those in Table 1. The clusters are numbered from I to V.

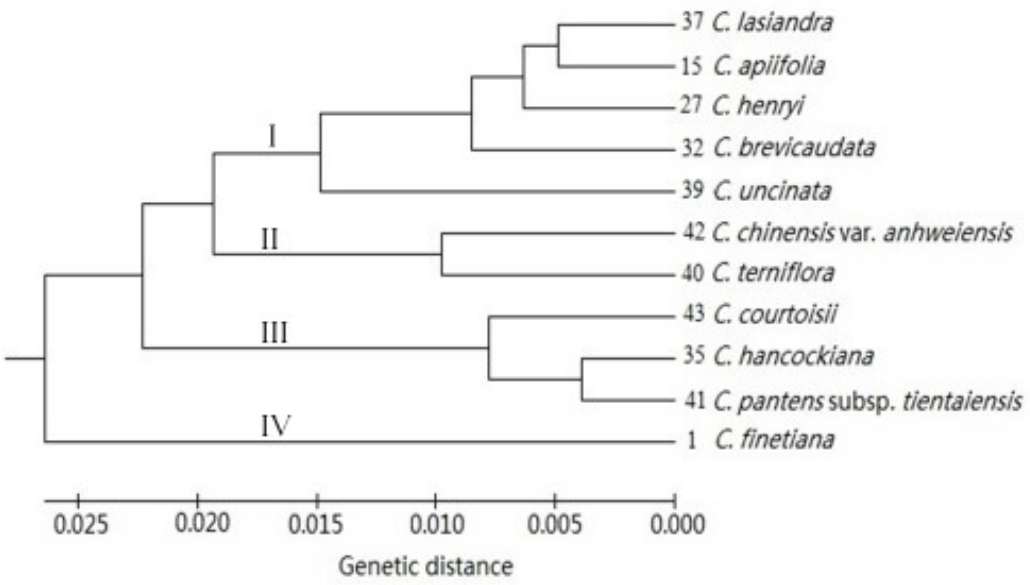

Fig. 2. A dendrogram of 11 *Clematis* species based on ITS sequence information. The sample numbers follow those in Table 1. The clusters are numbered from I to IV.